

WHAT IS CLAIMED IS:

1. A composition comprising a protein, wherein the protein comprises at least one unnatural amino acid and at least one post-translational modification, wherein the at least one post-translational modification comprises attachment of a molecule comprising a second reactive group by a [3+2] cycloaddition to the at least one unnatural amino acid comprising a first reactive group.

5 2. The composition of claim 1, wherein the molecule is a dye, a polymer, a derivative of polyethylene glycol, a photocrosslinker, a cytotoxic compound, an affinity label, a derivative of biotin, a resin, a second protein or polypeptide, a metal chelator, a 10 cofactor, a fatty acid, a carbohydrate, or a polynucleotide.

3. The composition of claim 1, wherein the first reactive group is an alkynyl or azido moiety and the second reactive group is an azido or alkynyl moiety.

4. The composition of claim 3, wherein first reactive group is the alkynyl moiety and the second reactive group is the azido moiety.

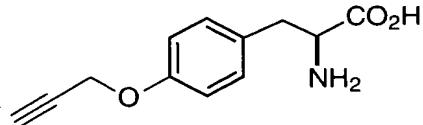
15 5. The composition of claim 4, wherein the unnatural amino acid comprises a *p*-propargyloxyphenylalanine.

6. The composition of claim 3, wherein the first reactive group is the azido moiety and the second reactive group is the alkynyl moiety.

7. The composition of claim 6, wherein the unnatural amino acid comprises a *p*-azido-L-phenylalanine.

8. The composition of claim 6, wherein the at least one post-translational modification is made *in vivo* in a eukaryotic cell.

9. A composition comprising an unnatural amino acid having the chemical structure:



25

10. The composition of claim 9, further comprising an orthogonal tRNA.

11. The composition of claim 10, wherein the unnatural amino acid is covalently bonded to the orthogonal tRNA.

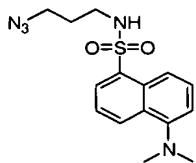
12. The composition of claim 10, wherein the unnatural amino acid is covalently bonded to the orthogonal tRNA though an amino-acyl bond.

5 13. The composition of claim 10, wherein the unnatural amino acid is covalently bonded to a 3'OH or a 2'OH of a terminal ribose sugar of the orthogonal tRNA.

14. A protein comprising the unnatural amino acid of claim 9.

15. A cell comprising the unnatural amino acid of claim 9.

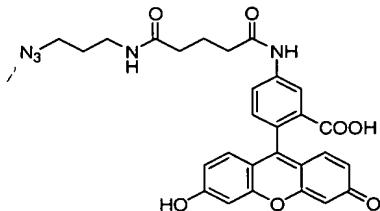
16. A composition comprising an azido dye having the structure:



10

4

17. A composition comprising an azido dye having the structure:



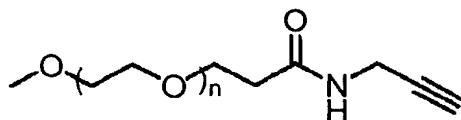
6

18. A protein comprising the azido dye of claim 16 or claim 17.

15 19. The protein of claim 18, further comprising at least one unnatural amino acid, wherein the azido dye is attached to the unnatural amino acid through a [3+2] cycloaddition.

20. The protein of claim 19, wherein the unnatural amino acid comprises an alkynyl amino acid.

20 21. A composition comprising an alkynyl polyethylene glycol having the structure:



wherein n is an integer between 100 and 2,000.

22. The composition of claim 21, wherein the alkynyl polyethylene glycol has a molecular weight of about 5,000 to about 100,000 Da.

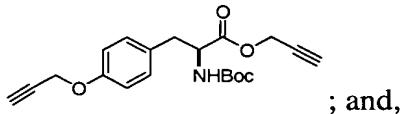
5 23. A protein comprising the alkynyl polyethylene glycol of claim 21.

24. The protein of claim 23, further comprising at least one unnatural amino acid, wherein the alkynyl polyethylene glycol is attached to an unnatural amino acid through a [3+2] cycloaddition.

10 25. The protein of claim 24, wherein the unnatural amino acid comprises an azido amino acid.

26. A method for synthesizing a p-(propargyloxy)phenylalanine compound, the method comprising:

suspending N-tert-butoxycarbonyl-tyrosine and  $K_2CO_3$  in anhydrous DMF; adding propargyl bromide to the reaction mixture of (a) and alkylating the hydroxyl and the 15 carboxyl group, resulting in an protected intermediate compound having the structure:



mixing the protected intermediate compound with anhydrous HCl in MeOH and deprotecting the amine moiety, thereby synthesizing the p-(propargyloxy)phenylalanine 20 compound.

27. The method of claim 26, further comprising:

dissolving the p-(propargyloxy)phenylalanine HCl in aqueous NaOH and MeOH and stirring at room temperature;

adjusting the pH of to pH 7; and,

25 precipitating the p-(propargyloxy)phenylalanine compound.

28. A method for synthesizing an azido dye, the method comprising:  
providing a dye compound comprising a sulfonyl halide moiety;  
warming the dye compound to room temperature in the presence of 3-  
azidopropylamine and triethylamine; and,  
5 coupling an amine moiety of the 3-azidopropylamine to the halide position of the  
dye compound, thereby synthesizing the azido dye.

29. The method of claim 28, wherein the dye compound comprises dansyl  
chloride, and wherein the azido dye comprises the composition of claim 16.

30. The method of claim 28, further comprising:  
10 purifying the azido dye from the reaction mixture.

31. A method for synthesizing an azido dye, the method comprising:  
providing an amine-containing dye compound;  
combining the amine-containing dye compound with a carbodiimide and 4-(3-  
azidopropylcarbamoyl)-butyric acid in a suitable solvent, and coupling a carbonyl group of  
15 the acid to the amine moiety of the dye compound, thereby synthesizing the azido dye.

32. The method of claim 31, wherein the carbodiimine comprises 1-ethyl-3-(3-  
dimethylaminopropyl) carbodiimide hydrochloride (EDCI).

33. The method of claim 31, wherein the amine-containing dye comprises  
fluoresceinamine, and the suitable solvent comprises pyridine.

20 34. The method of claim 31, wherein the amine-containing dye comprises  
fluoresceinamine and the azido dye comprises the composition of claim 17.

35. The method of claim 31, further comprising:  
precipitating the azido dye;  
washing the precipitate with HCl;  
25 dissolving the washed precipitate in EtOAc; and  
precipitating the azido dye in hexanes.

36. A method for synthesizing a propargyl amide polyethylene glycol, the  
method comprising: reacting propargylamine with polyethylene glycol (PEG)-

hydroxysuccinimide ester in an organic solvent at room temperature, resulting in the propargyl amide polyethylene glycol of claim 21.

37. The method of claim 36, wherein the organic solvent comprises CH<sub>2</sub>Cl<sub>2</sub>.

38. The method of claim 36, further comprising: precipitating the

5 propargylamide polyethylene glycol using ethyl acetate.

39. The method of claim 38, further comprising: recrystallizing the propargylamide polyethylene glycol in methanol; and drying the product under a vacuum.

40. A eukaryotic cell comprising an orthogonal aminoacyl-tRNA synthetase (O-RS), wherein the O-RS preferentially aminoacylates an orthogonal tRNA (O-tRNA) with at 10 least one unnatural amino acid in the eukaryotic cell, wherein:

(a.) the O-RS or a portion thereof is encoded by a polynucleotide sequence as set forth in any one of SEQ ID NO.: 20-25, a complementary polynucleotide sequence thereof, or a conservative variant thereof;

15 (b.) the O-RS comprises an amino acid sequence as set forth in any one of SEQ ID NO.: 48-63, or a conservative variant thereof;

(c.) the O-RS comprises an amino acid sequence that is at least 90% identical to that of a naturally occurring tyrosyl aminoacyl-tRNA synthetase (TyrRS) and comprises two or more amino acids selected from the group consisting of: glycine, serine, or alanine at a position corresponding to Tyr37 of *E. coli* TyrRS; aspartate at a position corresponding to 20 Asn126 of *E. coli* TyrRS; asparagine at a position corresponding to Asp182 of *E. coli* TyrRS; alanine, or valine, at a position corresponding to Phe183 of *E. coli* TyrRS; and, methionine, valine, cysteine, or threonine, at a position corresponding to Leu186 of *E. coli* TyrRS; or,

(d.) the O-RS aminoacylates the O-tRNA with the at least one unnatural amino acid 25 at least 50% as efficiently as does an O-RS having an amino acid sequence as set forth in SEQ ID NO.: 45.

41. The cell of claim 40, further comprising an orthogonal tRNA (O-tRNA), wherein the O-tRNA recognizes a selector codon and is preferentially aminoacylated with the at least one unnatural amino acid by the O-RS, wherein the O-tRNA is produced in a 30 cell by cellular processing of a nucleic acid corresponding to SEQ ID NO.:65, and the O-RS

comprises a polypeptide sequence selected from the group consisting of: SEQ ID NO.: 48-63, and a conservative variation thereof.

42. A polypeptide selected from the group consisting of:

(a) a polypeptide that comprises an amino acid sequence as shown in any one of

5 SEQ ID NO.: 48-63;

(b) a polypeptide that comprises an amino acid sequence encoded by a polynucleotide sequence as shown in any one of SEQ ID NO.: 20-35;

(c) a polypeptide that is specifically immunoreactive with an antibody specific for a polypeptide of (a), or (b);

10 (d) a polypeptide that comprises an amino acid sequence that is at least 90% identical to that of a naturally occurring tyrosyl aminoacyl-tRNA synthetase (TyrRS) and comprises two or more amino acids selected from the group consisting of: glycine, serine, or alanine at a position corresponding to Tyr37 of *E. coli* TyrRS; aspartate at a position corresponding to Asn126 of *E. coli* TyrRS; asparagine at a position corresponding to 15 Asp182 of *E. coli* TyrRS; alanine, or valine, at a position corresponding to Phe183 of *E. coli* TyrRS; and, methionine, valine, cysteine, or threonine, at a position corresponding to Leu186 of *E. coli* TyrRS;

20 (e) a polypeptide that comprises at least 20 contiguous amino acids of SEQ ID NO.: 36-48, or 86, and two or more amino acid substitutions selected from the group consisting of: glycine, serine, or alanine at a position corresponding to Tyr37 of *E. coli* TyrRS, aspartate at a position corresponding to Asn126 of *E. coli* TyrRS, asparagine at a position corresponding to Asp182 of *E. coli* TyrRS, alanine, or valine, at a position corresponding to Phe183 of *E. coli* TyrRS, and methionine, valine, cysteine, or threonine, at a position corresponding to Leu186 of *E. coli* TyrRS; and,

25 (f) an amino acid sequence comprising a conservative variation of (a), (b), (c), (d), or (e).

43. A composition comprising the polypeptide of claim 42 and an excipient.

44. An antibody or antisera specifically immunoreactive with the polypeptide of claim 42.

30 45. A composition comprising the polypeptide of claim 42 and an excipient.

46. An antibody or antisera specifically immunoreactive with the polypeptide of claim 42.
47. A polynucleotide selected from the group consisting of:
- (a) a polynucleotide comprising a nucleotide sequence as set forth in any one of SEQ ID NO.: 20-35;
  - (b) a polynucleotide that is complementary to or that encodes a polynucleotide sequence of (a);
  - (c) a polynucleotide encoding a polypeptide that comprises an amino acid sequence as set forth in any one of SEQ ID NO.: 48-63, or a conservative variation thereof;
- 10 (d) a polynucleotide that encodes a polypeptide of claim 42;
- (e) a nucleic acid that hybridizes to a polynucleotide of (a), (b), (c), or (d) under highly stringent conditions over substantially the entire length of the nucleic acid;
  - (f), a polynucleotide that encodes a polypeptide that comprises an amino acid sequence that is at least 90% identical to that of a naturally occurring tyrosyl aminoacyl-tRNA synthetase (*TyrRS*) and comprises two or more mutations selected from the group consisting of: glycine, serine, or alanine at a position corresponding to Tyr37 of *E. coli* *TyrRS*, aspartate at a position corresponding to Asn126 of *E. coli* *TyrRS*, asparagine at a position corresponding to Asp182 of *E. coli* *TyrRS*, alanine, or valine, at a position corresponding to Phe183 of *E. coli* *TyrRS*, and methionine, valine, cysteine, or threonine, 15 at a position corresponding to Leu186 of *E. coli* *TyrRS*;

20 (g) a polynucleotide that is at least 98% identical to a polynucleotide of (a), (b), (c), (d), (e), or (f); and,

    - (h) a polynucleotide comprising a conservative variation of (a), (b), (c), (d), (e), (f), or (g).

25 48. A vector comprising a polynucleotide of claim 47.

49. The vector of claim 48, wherein the vector comprises a plasmid, a cosmid, a phage, or a virus.

50. The vector of claim 48, wherein the vector is an expression vector.

51. A cell comprising the vector of claim 48.

52. A method of producing in a eukaryotic cell at least one protein comprising at least one unnatural amino acid, the method comprising:

growing, in an appropriate medium, a eukaryotic cell that comprises a nucleic acid that comprises at least one selector codon and encodes the protein; wherein the medium

5 comprises an unnatural amino acid and the eukaryotic cell comprises:

an orthogonal tRNA (O-tRNA) that functions in the cell and recognizes the selector codon; and,

an orthogonal aminoacyl tRNA synthetase (O-RS) that preferentially aminoacylates the O-tRNA with the unnatural amino acid, wherein the O-RS comprises an amino acid sequence that corresponds to SEQ ID NO.: 48-53.

53. A method of producing in a eukaryotic cell at least one protein comprising at least one unnatural amino acid, the method comprising:

growing, in an appropriate medium, a eukaryotic cell that comprises a nucleic acid that comprises at least one selector codon and encodes the protein; wherein the medium

15 comprises an unnatural amino acid and the eukaryotic cell comprises an orthogonal tRNA (O-tRNA) that functions in the cell and recognizes the selector codon and an orthogonal aminoacyl tRNA synthetase (O-RS) that preferentially aminoacylates the O-tRNA with the unnatural amino acid;

incorporating into the protein the unnatural amino acid in the eukaryotic cell,  
20 wherein the unnatural amino acid comprises a first reactive group; and

contacting the protein with a molecule that comprises a second reactive group; wherein the first reactive group reacts with the second reactive group to attach the molecule to the unnatural amino acid through a [3+2] cycloaddition.

54. The method of claim 53, wherein the molecule is a dye, a polymer, a derivative of polyethylene glycol, a photocrosslinker, a cytotoxic compound, an affinity label, a derivative of biotin, a resin, a second protein or polypeptide, a metal chelator, a cofactor, a fatty acid, a carbohydrate, or a polynucleotide.

55. The method of claim 53, wherein the first reactive group is an alkynyl or azido moiety and the second reactive group is an azido or alkynyl moiety.

56. The method of claim 55, wherein first reactive group is the alkynyl moiety and the second reactive group is the azido moiety.

57. The method of claim 56, wherein the unnatural amino acid comprises a *p*-propargyloxyphenylalanine.

5 58. The method of claim 55, wherein the first reactive group is the azido moiety and the second reactive group is the alkynyl moiety.

59. The method of claim 58, wherein the unnatural amino acid comprises a *p*-azido-L-phenylalanine.

60. A protein produced by the method of claim 53.

10 61. The protein of claim 60, wherein the protein is modified by at least one post-translational modification *in vivo* and wherein the post-translational modification is selected from the group consisting of: N-glycosylation, O-glycosylation, acetylation, acylation, lipid-modification, palmitoylation, palmitate addition, phosphorylation, and glycolipid-linkage modification.